Defence and Peace Economics

ISSN: 1024-2694 (Print) 1476-8267 (Online) Journal homepage: http://www.tandfonline.com/loi/gdpe20

Asymmetric information and third-party intervention in civil wars

J. Atsu Amegashie

To cite this article: J. Atsu Amegashie (2014) Asymmetric information and third-party intervention in civil wars, Defence and Peace Economics, 25:4, 381-400, DOI: 10.1080/10242694.2013.799935

To link to this article: http://dx.doi.org/10.1080/10242694.2013.799935

Published online: 20 Jun 2013.

Defence and Peace Economics, 2014
Vol. 25, No. 4, 381–400, http://dx.doi.org/10.1080/10242694.2013.799935

Routledge
Taylor & Francis Group

ASYMMETRIC INFORMATION AND THIRD-PARTY INTERVENTION IN CIVIL WARS

J. ATSU AMEGASHIE*

Department of Economics, University of Guelph, Guelph, Canada

(*Received 7 March 2012; in final form 9 April 2013*)

I study a two-period model of conflict with two combatants and a third party who is an ally of one of the combatants. The third party is fully informed about the type of her ally but not about the type of her ally's enemy. In a signaling game, I find that if the third party is unable to give a sufficiently high assistance to her ally, then there exists a unique separating equilibrium in which the third party's *expected* intervention causes her ally's enemy to exert more effort than in the absence of third-party intervention; this worsens the conflict.

Keywords: Perfect Bayesian equilibrium; Conflict; Intuitive criterion; Signaling; Third-party intervention

JEL Codes: D72, D74

1. INTRODUCTION

The literature on intrastate conflicts is large and growing. Within this literature, research on third-party intervention in intrastate conflicts has received very little attention from economists.[1]

In the post World War II era and the end of the cold war, there have been numerous third-party interventions in intrastate conflicts. These interventions have been in places such as Bosnia, Somalia, Haiti, the former Soviet republics, Vietnam, and Cambodia and have involved countries like Britain, China, France, the USA, and international organizations like the UN. Between 1944 and 1999, Regan (2002) identified 150 intrastate conflicts of which 101 had third-party interventions. He found that third-party interventions tend to worsen conflicts (see also Regan, 2000). Elbadawi and Sambanis (2000) also obtained a

*Email: jamegash@uoguelph.ca

[1]See Blattman and Miguel (2010) and Collier and Hoeffler (2007) for surveys of the literature on intrastate conflicts and Amegashie (2010) for a discussion of third-party intervention in conflicts.

similar result in their empirical work. And Diehl *et al*. (1996) made a similar claim in the case of UN interventions.[2]

Still on the preceding point, Lacina (2006) and Heger and Salehyan (2007) report that civil wars with interventions have significantly higher fatality levels than civil wars without interventions. Recent work by Cunningham (2010) suggests that external interventions significantly prolong conflicts. Of course, there is the identification problem of whether interventions make civil wars worse, or whether really bad civil wars are more likely to provoke intervention. As pointed out below, this explains why some scholars have used expected intervention to deal with the endogeneity of actual intervention.

On the preceding point, Gent (2008) argues that a selection effect explains his empirical finding that interventions on behalf of rebel groups are more effective than interventions on behalf of ruling governments. His game-theoretic model predicts that government-biased third parties intervene in tougher conflicts relative to rebel-biased interveners. Therefore, rebel-biased interventions are more effective than government-biased interventions.

Gent (2008) is one of several papers that studies biased interventions where third parties support one of the factions in a conflict. Using data from the International Crisis Behavior project, Carment and Harvey (2000) found that 140 out of 213 interventions in intrastate conflicts over the period 1918–1994 were clearly biased. In the postwar period, Regan (2000) also found that most interventions were biased. In his empirical work, Regan (2002) found that neutral interventions were less effective in ending conflicts than biased interventions. Betts (1996) argued that the idea of impartial intervention is a delusion and authors such as Watkins and Winters (1997) and Favretto (2009) have argued that biased interventions may be desirable.

While some of the aforementioned empirical studies have found that third-party intervention could worsen conflicts, exactly why this might be the case is not clear. This paper builds a model to explain how a third party's intervention in a conflict could worsen the conflict. There is a literature that emphasizes the moral hazard effects of third-party intervention in conflicts (e.g. Kuperman, 1996; Rowlands and Carment, 1998, 2006). This literature suggests that domestic groups which would not otherwise resort to political violence or escalate an ongoing conflict may be encouraged to do so by the prospect of outside support.[3] Explicit models of this argument include Rowlands and Carment (1998, 2006), Favretto (2009). However, they treat the size of the third party's military support as

[2]See also Balch-Lindsay *et al*. (2008). These papers used the duration of conflict to evaluate third-party interventions. Hence, third-party intervention worsens a conflict if it extends the duration of the conflict. In this paper, I use the effect of third-party intervention on the aggregate effort by the warring faction to evaluate the third party's intervention. This is reasonable because there is likely to be a positive correlation between the duration of conflict and the magnitude of the social costs in terms of the loss of life and property. In the related case of mediation, as opposed to military intervention, Regan and Stam (2000) undertook an empirical analysis of third-party intervention in conflicts and found that mediation in the earlier stages of a conflict are more effective and that late mediations worsen conflicts. Dixon (1996) found that mediation efforts and third-party activities to open or maintain lines of communication are the most effective to resolving conflicts.

[3]In a 31 May 2006 op-ed in the New York Times, Alan Kuperman claimed that intervention in Darfur was emboldening the rebels to fight on because the rebels who benefited from intervention rejected a proposed agreement. A similar argument has been made by some scholars with respect to the Kosovo crisis (see Grigorian, 2005). However, others do not think that the moral hazard argument is satisfactory (e.g. Crawford, 2005; Grigorian, 2005).

exogenous.[4] Other scholars argue that a lack of resolve and credibility within coalitions over the use of force create incentives for the escalation of conflicts (Diehl *et al.*, 1996; Harvey, 1998; Regan, 1996; Walter and Snyder, 1999).

In this paper, I focus on biased interventions. I study a conflict with incomplete information and two combatants who fight over two periods. One of the combatants has an ally (i. e. a third party) who wants to assist him with military support in the conflict. The third party is fully informed about the type of her ally but not about the type of her ally's enemy. The third party's ally is fully informed about the type of his enemy. There is a signaling game with one receiver (i.e. the third party) and two senders (i.e. the two combatants). I focus on the separating equilibrium of the game and adapt the intuitive criterion to restrict beliefs. I find that if the third party is unable to give a sufficiently high military assistance to her ally, then there is a perfect Bayesian separating equilibrium in which the third party's *expected* intervention worsens the conflict by inducing the enemy of the third party's ally to overinvest in arms in order to discourage the third party from helping her ally or to back off entirely from intervening in the conflict in period 2. Hence, the enemy of the third party's ally displays some bravado (i.e. overinvests in arms). Not only does third-party intervention lead to an increase in the effort of the ally's enemy, it also leads to an increase in the *aggregate* effort in the conflict.

The paper also shows that a third-party may rationally mistrust her ally by ignoring the ally's private and valuable information. Such mistrust may exist between allies who could be described as strange bedfellows. An example of this mistrust is the occasional claims by the USA that its allies (i.e. the governments of Pakistan, Iraq, and Afghanistan) are not doing enough to rein in their enemies in their various countries.

The result that the third party's intervention may worsen the conflict if the size of intervention is not big enough is consistent with the literature which argues that lack of resolve within coalitions over the use of force create incentives for the escalation of conflicts.

To deal with the endogeneity of third-party intervention, Elbadawi and Sambanis (2000) used *expected* intervention instead of actual intervention as the regressor in their empirical work. They found that *expected* third-party interventions worsen conflicts. Akcinaroglu and Radziszewski (2005) also reach a similar conclusion in their empirical work. These empirical papers lend some support to the theoretical result that *expected* intervention may worsen a conflict. Furthermore, as Wagner (2007, 229) observed: 'expectations about possible interventions may play a role in motivating an internal conflict even if outsiders never intervene in it.'

The result that the *expectation* of third-party support could increase the intensity of conflict, at least, for the ally's enemy has implications for empirical work. It suggests that conflicts may worsen prior to a publicly known and biased intervention and may improve after the intervention. Hence, part of the subsequent reduction in the intensity of the conflict is not necessarily due to actual intervention *per se*. Therefore, while a biased intervention by a third party may have a positive effect on a conflict, this effect may be

[4]There is a wide literature on third-party intervention which uses the Fearon (1994, 1995) crisis-bargaining model. Favretto (2009) uses a model of incomplete information in the framework of Fearon's (1994, 1995) crisis-bargaining model. In contrast to the present model, she assumes exogenous probabilities of success for the warring factions. Using a crisis-bargaining model with no signaling, Werner (2000) considers a model of incomplete information to determine a third party's incentives to intervene in a conflict. She finds that an attacker can manipulate the stakes of war by making it low enough that, for a third party, the benefits of intervention do not justify the costs (see also Yuen, 2009). In my model, no one can manipulate the stakes of the conflict. And in Werner (2000), the threat of intervention does not worsen the conflict.

overstated in empirical work (i.e. the relevant coefficient in regressions may be biased upwards).[5]

In my two-period model, the intensity of conflict is increased not only because of the *expectation* of the third party's intervention as is the case in period 1, but because the third party's *actual* intervention reduces the ally's marginal cost of conflict inducing the ally to be more aggressive in period 2 than he otherwise would have been. This is consistent with the moral hazard literature of *actual* intervention mentioned above. However, this effect, while present in my model, is not the focus of this paper.

It is important to note that although the third party's ally (i.e. faction A) knows her enemy's (i.e. faction B) type, she cannot communicate this information to the third party. A reason is that faction A's information is 'soft information' not 'hard information.' But more importantly, given this 'soft information,' the third party cannot believe whatever faction A says because faction A has the incentive to lie about faction B's type when faction B is strong.

The main point of this paper is *not* to make the *general* point that signaling can lead to undesirable outcomes. A contribution of this paper is to theoretically explain how *expected* third-party intervention could worsen conflicts. Furthermore, as discussed above, other interesting insights such as the implications of the results for econometric work and the conditions under which third-party intervention may worsen conflicts emerge from the analysis. In Section 3, I discuss alternative models and compare them with my model.

The paper is organized as follows. Section 2 presents a model and analyses of signaling in a conflict. Section 3 discusses issues of robustness and Section 4 concludes the paper.

2. A MODEL OF THIRD-PARTY INTERVENTION AND SIGNALING

I use a well-known model of conflict and contests (e.g. see Epstein and Gang, 2009; Konrad, 2009) which assumes that the combatants cannot commit to a peaceful resolution through bargaining. Therefore, costly conflict is inevitable. This model is different from crisis-bargaining models pioneered by Fearon (1994, 1995).[6]

Consider two risk-neutral factions, A and B, in a conflict over a region (country). Faction A is an incumbent who governs the country and faction B is a challenger to faction A's rule. There is a risk-neutral third party, C, who is an ally of faction A. There are two time periods, 1 and 2. In each period, there could be a battle (conflict) between A and B. Faction A's valuation of controlling a proportion, $P_A^j \in [0,1]$, of the land (country) in period j is $P_A^j\ V \geqslant 0$, where $V > 0$, $j = 1$, 2. Faction B's corresponding valuation is P_B^j W_H with probability $q \in (0,1)$ and $P_B^j\ W_L$ with probability $1-q$, where $W_H > W_L > 0$ and

[5]Notice that this argument is different from the argument in Elbadawi and Sambanis (2000) and Akcinaroglu and Radziszewski (2005) for using expected intervention as an explanatory variable in a regression. Unlike these authors, I am not arguing that expected intervention should be used as an instrumental variable in order to deal with the endogeneity of actual intervention. I am arguing that expected intervention and actual intervention may be distinct explanatory variables in a regression that seek to explain the intensity of conflict. Still, their finding suggests that expected intervention could have a negative effect on conflicts.

[6]In a seminal two-player crisis-bargaining model in the shadow of conflict, Fearon (1994, 1995) used costly signaling as an explanation of war. In these models, the threat of conflict, not necessarily actual conflict, is used to make credible demands in crisis bargaining. In my model, conflict is inevitable and I do not allow bargaining between the third party and his ally on one hand and the ally's enemy on the other. Crisis bargaining models are typically used to explain the conditions under which conflict will occur and the role of private information is at the heart of this literature (see Walter, 2009, for a review). I am interested in the conditions under which third-party intervention will escalate an ongoing conflict. Also, signaling in crisis bargaining models is different from signaling in the classical sense.

$P_A^j = 1 - P_B^j$. If faction A controls a proportion P_A^j of the land (country), the third party's valuation is $P_A^j\ S > 0$ where $S > 0$. Since a player with a higher valuation in contest is the same as a player with a lower unit cost of exerting effort (e.g. see Clark and Riis, 1998), I shall refer to the high-valuation type of faction B as the strong type and the low-valuation type as the weak type. Because of risk-neutrality, the proportions could alternatively be interpreted as probabilities of victory in the conflict with S, V, and W_k as the players' valuations of victory, k = H, L. Note also that the third party's valuation, S, is a measure of his strength or military capability and, as will be obvious in the subsequent analysis, the third party will intervene in the conflict if his valuation is sufficiently high.

In *each* period, the factions fight over the *same* piece of land or resource. They invest in a composite military good (hereafter referred to as arms, armed investments, or simply investment) that could be thought of being made up of weapons and soldiers. In period 1, I denote the factions' investments in arms by G and in period 2, I denote it by X. So, for example in period 1, factions A and B invest in G_A and G_B units of arms and control the proportions $P_A^1 = G_A/(G_A + G_B)$ and $P_B^1 = 1 - P_A^1$ of the land (country). A similar function describes the proportions in period 2. Throughout the analysis, G_A, G_B, X_A, and X_B are assumed to be nonnegative.

I assume that the third party can help faction A in period 2 but not in period 1. For example, in the case of the USA, this may be due to delays in congressional approval of funding for military support of her allies. This assumption is not crucial. What I need is that in period 1, the third party is not fully informed about the type of her ally's enemy (i.e. faction B) and her assistance decision in period 1 is taken before faction B moves.[7] Therefore, the third party could already have intervened in the conflict in period 1 at some level of exogenous military assistance that I have set to zero.[8]

In the absence of third-party intervention, each faction has a unit cost of arms equal to 1. Intervention can take various forms and can be modeled in different ways. I follow the formulation in Chang *et al*. (2007). In particular, when the third party spends M dollars on military assistance to faction A, it affects faction A's unit cost of arms through some reduced-form relationship such that faction A's unit cost of arms decreases from 1 to $1/(1+M)^{\theta}$, where θ measures the degree of effectiveness with which a dollar of assistance reduces faction A's unit cost of arming and $0 < \theta < 1$. I discuss the implications of this formulation of third-party intervention in Section 3.1.

The timing of actions is as follows:

Period 1 (or the first battle):

Stage 1: Nature chooses faction B's type (valuation): W_H or W_L. This becomes common knowledge to factions A and B but not to the third party. The third party only knows that $\Pr(W_H) = q \in (0,1)$. The valuation, V, of faction A is common knowledge.
Stage 2: Faction B chooses his investment in arms.
Stage 3: Faction A observes B's choice and chooses his investment in arms.

[7]This means that I could demonstrate my result in a one-period model with more stages. However, the two-period model below is more convenient for exposition.

[8]In the model, signaling occurs in period 1. Imagine that there is an *exogenous* third-party military assistance in period 1 but no intervention in period 2. That is, the third party has decided to *exogenously withdraw* from the conflict in period 2. Call this the no-intervention case. In this case, there will be no signaling in period 1. The equilibrium of the game in period 1 when the third party has incomplete information is the same as the equilibrium when the third party has complete information. Now suppose in period 2, the third party decides to withdraw or stay in the conflict based on the outcome of the conflict in period 1. Then, the analysis in the paper goes through.

Period 2 (or the second battle):

Faction B's valuation in period 1 is also his valuation in period 2, but this is still only known by factions A and B and may remain unknown to the third party.[9] Faction A's valuation is still common knowledge.

Stage 1: The third party chooses how much help to offer faction A.
Stage 2: Faction A observes the third party's choice and chooses his investment in arms.
Stage 3: Faction B observes the choices of the third-party and faction A and chooses his investment in arms.

The timing of moves between factions A and B in period 1 is not the same as in period 2. This may come across as odd but it should not.[10] In a situation of war, it is not unreasonable to believe that given that B moved first in the current battle (i.e. conflict in period 1) faction A – perhaps to avoid another surprise attack or to exact revenge and also fight for the land – is likely to move first in the next battle (i.e. conflict in period 2). In fact, in a situation of war, there is no presumption that a faction which launched an offensive attack in one battle will necessarily be the first to launch an offensive attack in the next battle. Furthermore, the desire to exact revenge while at the same time trying to gain control over an asset (e.g. land) is consistent with the aforementioned sequence of moves. The faction exacting revenge (call it faction A) in the current battle will be the first mover and the faction (i.e. B) that responds to this attack will be the second mover. This is not inconsistent with faction B having been the first mover in the previous battle. This may be precisely what faction A may try to avoid by moving first in the current battle.[11]

2.1. Equilibrium in Period 2

The game described above is a signaling game with two senders, A and B, and one receiver (i.e. the third party). However, as I show below, signaling will occur in period 1 but not in period 2. I look for perfect Bayesian Nash equilibria of this game and restrict attention to separating equilibria in pure strategies. I adapt the intuitive criterion to restrict beliefs.

Consider period 2. Given that in period 2, factions A and B engage in conflict after the third party has given her assistance and that this assistance cannot be withdrawn, it follows that the game between A and B in period 2 will involve no signaling and it is, therefore, the same as the complete-information game analyzed in Chang *et al*. (2007).[12]

[9]The assumption that faction B's valuation in period 2 is the same as his valuation in period 1 is crucial. If nature were to move again in period 2, there will be no need for signaling in period 1.

[10]In addition to the reasons given below, note that in the influential alternating-offers bargaining game pioneered by Ariel Rubinstein, the timing of moves between the two bargainers is not the same in each period.

[11]For a paper that examines revenge in conflicts, see Amegashie and Runkel (2012).

[12]Chang *et al*. (2007) and Chang and Sanders (2009) used a very standard model of sequential contests. Their contribution was the introduction of third-party intervention into the Tullock contest-type models of Grossman and Kim (1995), Gershenson and Grossman (2000), Leininger (1993), and Morgan (2003). See Konrad (2009) for a review of sequential contests. Assuming complete information and simultaneous moves, this model of conflict has also been used to analyze third-party intervention by Amegashie and Kutsoati (2007) who endogenized a third party's choice of her ally while Carment and Rowlands (1998), Rowlands and Carment (2006), and Siqueira (2003) took the third party's ally as given and examined the effect of third party's intervention on conflicts. Furthermore, there is a small literature on signaling in conflicts and contests. However, this literature considers only two players. It does not consider a third party or third-party intervention and so its focus is entirely different from this paper.

In stage 3 of period 2, faction B of type k chooses $X_{k|\mathrm{B}}$ to maximize

$$\Pi^2_{k|\mathrm{B}} = \frac{X_{k|\mathrm{B}}}{X_{k|\mathrm{B}} + X_{k|\mathrm{A}}} W_k - X_{k|\mathrm{B}}, \tag{1}$$

where $X_{k|\mathrm{A}}$ is the armed investment of faction A when his opponent is faction B of type k, $k = \mathrm{H}$, L.

Then, the following Kuhn–Tucker condition must hold:

$$\frac{\partial \Pi^2_{k|\mathrm{B}}}{\partial X_{k|\mathrm{B}}} = \frac{X_{k|\mathrm{A}}}{\left(X_{k|\mathrm{A}} + X_{k|\mathrm{B}}\right)^2} W_k - 1 \le 0; X_{k|\mathrm{B}} = 0 \text{ if} \frac{\partial \Pi^2_{k|\mathrm{B}}}{\partial X_{k|\mathrm{B}}} < 0. \tag{2}$$

The condition in (2) implies that the best response function for B is

$$X_{k|\mathrm{B}} = \max\left[0, \sqrt{W_k X_{k|\mathrm{A}}} - X_{k|\mathrm{A}}\right] \tag{3}$$

Therefore, $X_{k|\mathrm{B}} = 0$ if $X_{k|\mathrm{A}} \ge W_k$. In this case, faction B will not challenge faction A.

In stage 2 of period 2, faction A facing faction B of type k chooses $X_{k|\mathrm{A}}$ to maximize

$$\Pi^2_{k|\mathrm{A}} = \frac{X_{k|\mathrm{A}}}{X_{k|\mathrm{B}} + X_{k|\mathrm{A}}} V - \frac{1}{(1+M)^\theta} X_{k|\mathrm{A}}, \tag{4}$$

where M is the third party's assistance to faction A. As will be shown shortly, this will depend on the third party's belief that faction B is strong (weak).

From (3), put $X_{k|\mathrm{B}} = \sqrt{W_k X_{k|\mathrm{A}}} - X_{k|\mathrm{A}}$ into (4) and simplify to get

$$\Pi^2_{k|\mathrm{A}} = V\sqrt{\frac{X_{k|\mathrm{A}}}{W_k}} - \frac{1}{(1+M)^\theta} X_{k|\mathrm{A}} \tag{4a}$$

Then, the unique optimal investment in arms by faction A noting that $X_{k|\mathrm{B}} = 0$ if $X_{k|\mathrm{A}} \ge W_k$ is

$$\hat{X}_{k|\mathrm{A}} = \min\left[W_k, \frac{V^2(1+M)^{2\theta}}{4W_k}\right] \tag{5}$$

Then, $\hat{X}_{k|\mathrm{B}} = \sqrt{W_k \hat{X}_{k|\mathrm{A}}} - \hat{X}_{k|\mathrm{A}}$ which gives faction B of type k's unique armed investment as

$$\hat{X}_{k|\mathrm{B}} = \max\left[0, \frac{V(1+M)^\theta}{2} - \frac{V^2(1+M)^{2\theta}}{4W_k}\right] \tag{6}$$

Suppose that $M = 0$. Then, using Equation (6), $\hat{X}_{k|\mathrm{B}} > 0$ if

$$V\left[\frac{1}{2} - \frac{V}{4W_k}\right] > 0 \Rightarrow V < 2W_k, \tag{7}$$

$k = \text{H, L}$. I assume that (7) holds. Hence, faction A cannot deter faction B without the assistance of the third party.[13] This also means, from (5), that

$$\hat{X}_{k|\mathrm{A}} = V^2(1+M)^{2\theta}/4W_k. \tag{8}$$

The equilibrium proportion of the land controlled by faction A, whose opponent is faction B of type k, is

$$\hat{P}^2_{k|\mathrm{A}} = \frac{\hat{X}_{k|\mathrm{A}}}{\hat{X}_{k|\mathrm{A}} + \hat{X}_{k|\mathrm{B}}} = \frac{V(1+M)^\theta}{2W_k}, \tag{9}$$

where $V(1+M)^\theta \leq 2W_\mathrm{L}$.[14]

The proportion of the land to faction B of type k in period 2 is

$$\hat{P}^2_{k|\mathrm{B}} = 1 - \frac{(1+M)^\theta V}{2W_k}, \tag{10}$$

$k = \text{H, L}$.

Faction B of type k's equilibrium payoff in period 2 is

$$\hat{\Pi}^2_{k|\mathrm{B}} = \frac{(W_k - 0.5(1+M)^\theta V)^2}{W_k}, \tag{11}$$

and faction A's payoff is

$$\hat{\Pi}^2_{k|\mathrm{A}} = \frac{V^2(1+M)^\theta}{4W_k}, \tag{12}$$

$k = \text{H, L}$.

Clearly, faction B's payoff in (10a) is decreasing in the third party's military assistance to faction A while faction A's payoff in (10b) is increasing in the third party's assistance.

Now consider the third party's problem in stage 1 of period 2. Given the observed investments in arms by factions A and B in period 1, let $\mu \in [0,1]$ be the third party's belief in period 2 that faction B is a strong type. Then, the third party will choose her assistance M to faction A to maximize

$$\Omega^2_\mathrm{C}(\mu) = (1-\mu)\hat{P}^2_{\mathrm{L|A}}S + \mu\hat{P}^2_{\mathrm{H|A}}S - M. \tag{13}$$

Like most formal models of third-party intervention, the third party does not intervene because she cares about the welfare of her ally. She only cares about her own material welfare. However, she and her ally benefit more from a higher probability of victory of her ally or control over a bigger proportion of the land by her ally.

[13] The parameters of my numerical example are such that this feature of the model is preserved. If $V \geqslant 2W_k$ for all k, then there is no conflict even if the third party does not intervene, $k = \text{H, L}$. In this case, third-party intervention is not necessary, which is not a desirable feature of a model of third-party intervention. In this equilibrium, faction A is sufficiently armed (including the number of soldiers) leading faction B to acquiesce resulting in no conflict. See Grossman and Kim (1995) and Gershenson and Grossman (2000) for a discussion of this equilibrium. This equilibrium is not possible if factions A and B move simultaneously.

[14] This means that the third party's military assistance does not exceed what is required to deter the weak type of faction B. It implies that the strong type of faction B is not deterred in spite of the third party's assistance. I construct an example that satisfies this condition.

Putting (8) into (11) and maximizing gives the third party's optimal military assistance as:

$$\hat{M}(\mu) = \left[\frac{\theta SV}{2}\left(\frac{\mu}{W_{\mathrm{H}}} + \frac{1-\mu}{W_{\mathrm{L}}}\right)\right]^{1/(1-\theta)} - 1. \tag{14}$$

I assume that $\hat{M}(\mu) > 0 \forall \mu \in [0, 1)$ but $\hat{M}(\mu) \geq 0$ for $\mu = 1$.

Given (12) and $W_{\mathrm{H}} > W_{\mathrm{L}}$, it is obvious that the third party's military assistance is decreasing in her belief that her ally's enemy is a strong type. The intuition is that the higher is the third party's belief that faction B is strong, the smaller is the marginal return to her military assistance.[15] Therefore, this result and Equation (10) imply that faction B's payoff in period 2 is increasing in the third party's belief that he is strong. This explains why faction B may overinvest in arms in period 1. This is costly to him in period 1 but beneficial in period 2 because it will cause the third party to reduce her assistance to faction A.

2.2. Separating Equilibrium in Period 1

It is important to note that if faction A is choosing G_{A} to maximize only his payoff in period 1, then he will choose $G_{\mathrm{A}} = \max\left[0, \sqrt{VG_{\mathrm{B}}} - G_{\mathrm{B}}\right]$.

Remark: For want of a better expression, I say that faction A invests in signaling if his investment conveys some information to the third party about faction B's type. Otherwise, faction A does not invest in signaling. Given that the third party's choice of military support is influenced by her belief that faction B is strong and given that her military support occurs in period 2, faction A invests in signaling if and only if when choosing his investment (effort) in period 1, he includes his payoff in period 2 in his optimization problem. Hence, faction A does not invest in signaling if in choosing his investment in period 1, he maximizes *only* his payoff in period 1. In particular, faction A does not signal if, for a given G_{B}, he chooses $G_{\mathrm{A}} = \max\left[0, \sqrt{VG_{\mathrm{B}}} - G_{\mathrm{B}}\right]$, where $G_{\mathrm{A}} = 0$ if $G_{\mathrm{B}} \geqslant V$. It is important to note that this argument is also driven by fact that $G_{\mathrm{A}} = \max\left[0, \sqrt{VG_{\mathrm{B}}} - G_{\mathrm{B}}\right]$ is independent of faction B's type. However, by 'no signaling,' I do not mean that faction A is not informed about B's type, but rather that A's strategy in period 1 is not distorted by a concern to influence the third party's beliefs.

While it is conceivable that, given some beliefs by the third party, $G_{\mathrm{A}} = \max\left[0, \sqrt{VG_{\mathrm{B}}} - G_{\mathrm{B}}\right]$ might be an investment level that maximizes the sum of faction A's payoff over periods 1 and 2, it is reasonable to assume that the third party interprets this level of investment as no signaling by faction A because, given G_{B}, this is what maximizes faction A's period-1 payoff regardless of the third party's beliefs.

Let $G^*_{k|\mathrm{B}}$ be the investment of faction B of type k in period 1 in the benchmark case of full information, $k = \mathrm{H}, \mathrm{L}$. This is the full-information level of investment.

Lemma 1: There is a unique separating equilibrium in period 1 in which the strong type of faction B overinvests by choosing $\hat{G}_{\mathrm{H|B}} > G^*_{\mathrm{H|B}}$ and the weak type chooses $G^*_{\mathrm{L|B}}$. Faction A chooses $G^*_{\mathrm{L|A}} = \sqrt{VG^*_{\mathrm{L|B}}} - G^*_{\mathrm{L|B}}$ if faction B is weak and $\hat{G}_{\mathrm{H|A}} = \sqrt{V\hat{G}_{\mathrm{H|B}}} - \hat{G}_{\mathrm{H|B}}$ if faction B is strong.[16] The third party's equilibrium beliefs in period 2 are $\mu(G^*_{\mathrm{L|B}}, G^*_{\mathrm{L|A}})$

[15]This can easily be seen by putting (8) into (11).

[16]Of course, if the strong type of faction B's investment $\hat{G}_{\mathrm{H|B}} \geqslant V$, then faction A will choose a zero effort in period 1. The reader should be able to verify that this does not affect any of the analyses. However, in the subsequent numerical example, I choose parameters to ensure that faction A's equilibrium effort in period 1 is positive.

$=0$ and $\mu(\hat{G}_{\text{H|B}},\ G^*_{\text{H|A}}) = 1$. Then using Equation (12), the third party's military assistance to faction A in period 2 when faction B is strong is $\hat{M}(1)$ and her assistance when faction B is weak is $\hat{M}(0)$.

Proof: see appendix A.

Suppose that the third party's valuation, $S>0$, is so high that $\hat{M}(1)$ is such that $W_{\text{H}} - 0.5(1+\hat{M}(1))^{\theta}V \leq 0$. Then, $W_{\text{L}} - 0.5(1+\hat{M}(0))^{\theta}V<0$ because $\hat{M}(0)>\hat{M}(1)$ and $W_{\text{H}}>W_{\text{L}}$. Then, the third party will give enough military support to faction A to deter faction B in period 2 regardless of faction B's type. Then, faction B gains nothing by signaling in period 2, so he will not display any bravado. Therefore, faction B will only display bravado if the third party will *not* give a sufficiently high military assistance to faction A given that she believes that faction B is strong: $\hat{M}(1)<(2W_{\text{H}}/V)^{\theta} - 1$.[17]

This gives the following proposition:

Proposition 1: *If (i) the third party is not willing to intervene in a big way (i.e. $\hat{M}(1)<(2W_{\text{H}}/V)^{\theta} - 1$),[18] (ii) the solution to Equation (A.3) is such that $\hat{G}_{\text{H|B}}>G^*_{\text{H|B}}$, and (iii) the inequalities in (A.6) and (A.7) hold, then there exists a unique[19] perfect Bayesian separating equilibrium with bravado that satisfies an adapted intuitive criterion. In this equilibrium, the strong type of faction B is too aggressive by choosing a unique armed investment that is higher than his full-information level of investment in order to induce the third party to reduce his military assistance to faction A (i.e. the third party's ally) or back off from intervening in the conflict. The weak type of faction B chooses his unique full-information investment.*

2.3. Aggregate Destruction or Fatalities in the Conflict

As a proxy for aggregate fatalities or destruction in the conflict, I use the aggregate effort (investment) in the conflict. Using aggregate destruction in the conflict is consistent with the criterion used in empirical works, discussed in Section 1, which claim that third-party intervention worsens conflicts.

Suppose the conditions in proposition 1 hold. When faction B is strong, the aggregate effort (investment) in period 1 in the separating equilibrium is $\hat{G}_{\text{H|A}} + \hat{G}_{\text{H|B}} = \sqrt{V\hat{G}_{\text{H|B}}}>\sqrt{VG^*_{\text{H|B}}}$. Therefore, intervention by the third party increases aggregate effort in period 1. This is the moral hazard effect of third-party intervention which stems from *expected* intervention. If faction B is weak, aggregate effort in period 1 is $G^*_{\text{L|A}} + G^*_{\text{L|B}} = \sqrt{VG^*_{\text{L|B}}}$.

In period 2, the aggregate effort (investment) is $\hat{X}_{k|\text{A}} + \hat{X}_{k|\text{B}} = \sqrt{W_k\hat{X}_{k|\text{A}}}$, where $\hat{X}_{k|\text{A}}$ is given by Equation (5a). Then, given that $\hat{X}_{k|\text{A}}$ is increasing in the third party's military

[17] In other cases, the requirement to keep civilian causalities at a minimum means that a third party and her ally cannot deploy their full and combined military might.

[18] For example, if $\theta=0.5$, this condition is $8(W_{\text{H}})^2 - SV^2>0$.

[19] See appendix B for a proof that there is no pooling equilibrium. However, proving the nonexistence of a pooling equilibrium is not crucial. Focusing on separating equilibria is sufficient for my purposes because my goal is to construct an equilibrium in which third-party intervention may worsen a conflict. Therefore, what matters is to show that such an equilibrium exists. But, the nonexistence of a pooling equilibrium strengthens the results because it gives a stronger comparative static prediction.

assistance regardless of faction B's type, it follows that the third party's intervention increases the aggregate effort in period 2. This is the moral hazard effect of third-party intervention which stems from *actual* intervention (discussed in Section 1).

These results give the following proposition:

Proposition 2: *Suppose the conditions in proposition 1 hold, then third-party intervention may worsen the conflict. In particular, in period 1, the third party's expected intervention worsens the conflict when faction B is strong but has a neutral effect when faction B is weak. The third party's actual intervention worsens the conflict in period 2 regardless of faction B's type*.

In the separating equilibrium, the strong type of faction B displays bravado because of the third party's limited information about faction B. Furthermore, the third party can nullify this behavior if she can commit to a given level of assistance based on her prior beliefs. This can be summarized in the following corollary:

Corollary 1: *The expectation of a third party's assistance to an ally coupled with the third party's limited information about the strength of the enemy of her ally can be strategically exploited by the enemy. However, the ability of the enemy to strategically gain from his superior information no longer exists if the third party can commit to a given level of assistance based on her prior beliefs*.

3. FURTHER DISCUSSION AND ROBUSTNESS OF RESULTS

3.1. On the Effects of and Nature of Third-Party Intervention

As shown above, the third party's actual intervention in period 2 makes her ally too aggressive and thus escalates the conflict. This effect of actual intervention is not the focus of this paper. The focus of this paper is on expected intervention.

In this paper, a third party's *expected* intervention causes her ally's *enemy* to be more aggressive. However, there are instances where it can be plausibly argued that a third party's expected intervention can cause her ally to be more aggressive. For example, rebel groups in Libya may have fought harder in order to convince Western powers that providing them air support may create a significant chance of Gaddafi's removal. Also, after the 1991 Kuwait War, the Kurds and Shias rose up against Saddam Hussein expecting American military support.

However, the effects of expected intervention in the preceding paragraph cannot explain the argument by some scholars that a lack of resolve and credibility within coalitions over the use of force create incentives for the escalation of conflicts (Diehl *et al*., 1996; Harvey, 1998; Regan, 1996; Walter and Snyder, 1999). In fact, the effect of expected intervention in the preceding paragraph is likely to arise for precisely the opposite reason: the expectation by the *ally* that the third party has a strong resolve and so will intervene in a big way. As indicated above in the case of rebel groups who fought Gaddafi in Libya, if the ally expects the third party to intervene in a big way, he may fight harder in order to convince the third party to intervene.

The present model yields an effect of expected intervention that crucially depends on whether the third party intervenes in a big way and it is driven by the reaction of *ally's enemy*. Unlike the effect in the preceding paragraph, the ally's enemy (not the ally) fights harder if the third party is not expected to intervene in a big way and does not change his behavior if the third party intervenes in a big way. The analyses also show that a third party may ignore the signals by an informed ally, a result that cannot be obtained in a model which predicts that a third party's expected intervention will make his ally more aggressive. It is pointless for the ally to use extra aggression to signal to the third party if these signals will be ignored.

One may argue that the third party (e.g. the USA) may be directly involved in the conflict by committing troops and so military assistance need not take the form of reducing the ally's cost of conflict. What is important for my results is not whether the third party is directly involved in the conflict. What I need is that the third party's assistance or effort is decreasing in her belief that the enemy is strong. Moreover, third parties do intervene in conflicts by giving financial support for military expenditure without directly getting involved in the conflict (e.g. the USA's support of Angola's UNITA rebels during the cold war). Elbadawi and Sambanis (2000) define external intervention '… as a unilateral intervention by one (or more) third party government(s) in a civil war in the form of military, *economic* or mixed assistance in favor of either the government or the rebel movement involved in the civil war.'

There is nothing in the model which suggests that the third party's assistance is not an in-kind transfer. The third party's assistance could involve military training, the provision of intelligence information, military equipment, etc. These could have the effect of reducing his ally's cost of effort.

3.2. An Alternative Analysis of the Third Party's Behavior

In the present model, the third party's assistance is decreasing in her belief that her ally's enemy is strong. As explained earlier, this is because the return to her investment (assistance) is smaller when her ally's enemy is strong. However, another plausible scenario is that the third party is more likely to assist her ally when the ally's enemy is strong. For example, the benefits of defeating a strong enemy might be higher because there may otherwise be a higher probability of such an enemy challenging the third party's ally in the future if he is not defeated soundly in the current period. Or losing to a strong enemy makes it more difficult to fight back in the future. That is, if you lose to a strong enemy, you are less likely to get a second chance relative to losing to a weak enemy. In either case, the enemy has the incentive to feign weakness (i.e. under-invest in arms) rather than display bravado by overinvesting in arms.

It is possible to model how the aforementioned higher future benefit of defeating a stronger enemy may arise. However, I simply consider a reduced-form version of this argument by assuming that the third party's benefit when a strong enemy is defeated is S_{H} and when a weak enemy is defeated is S_{L}, where $S_{\mathrm{H}} \geq S_{\mathrm{L}} > 0$. Then, analogous to Equation (11), the third party will choose her assistance M to faction A to maximize

$$\tilde{\Omega}^2_{\mathrm{C}}(\mu) = (1-\mu)\frac{V}{2W_{\mathrm{L}}}(1+M)^{\theta}S_{\mathrm{L}} + \mu\frac{V}{2W_{\mathrm{H}}}(1+M)^{\theta}S_{\mathrm{H}} - M \tag{15}$$

The third party's optimal military assistance is:

$$\tilde{M}(\mu) = \left[\frac{\theta V}{2}\left(\mu\frac{S_{\mathrm{H}}}{W_{\mathrm{H}}} + (1-\mu)\frac{S_{\mathrm{L}}}{W_{\mathrm{L}}}\right)\right]^{1/(1-\theta)} - 1 \tag{16}$$

Then, $\partial\tilde{M}/\partial\mu > 0$ if $S_{\mathrm{H}}/W_{\mathrm{H}} > S_{\mathrm{L}}/W_{\mathrm{L}}$. Given that $W_{\mathrm{H}} > W_{\mathrm{L}}$, a necessary condition for this result is $S_{\mathrm{H}} > S_{\mathrm{L}}$. Clearly, if the benefit of defeating a strong enemy is higher than the benefit of defeating a weak enemy, then it is possible that the third party's assistance to her ally is increasing in her belief that her ally's enemy is strong. Then, we can construct an equilibrium in which the ally's enemy underinvests in arms in period 1. Also, when $S_H/W_H = S_L/W_L$, then $\partial\tilde{M}/\partial\mu = 0$. Then, there will be no signaling by either faction and third-party intervention will not worsen the conflict.

Of course, in this paper, I have assumed that $S_H = S_L = S$. Or more generally, I have assumed that $S_H/W_H < S_L/W_L$. In this case, equilibria with bravado exist and third-party intervention may worsen the conflict.

4. CONCLUSION

Expected third-party intervention may have a perverse effect on conflicts. This paper has shown that this may occur through a display of bravado by the enemy of the third party's ally. But in the separating equilibrium stated in proposition 1, bravado does not occur if the enemy of the third party's ally is weak.

When faction B is weak, the *expected* third party's intervention does not worsen the conflict in period 1. Hence, by using signaling as a possible mechanism through which a third party's intervention might lead to a perverse outcome, the paper sheds some light on the conditions under which *expected* third-party intervention may or may not worsen conflicts. Indeed, if the third party will not actually intervene in a big way, then it might be better if the third party does not intervene at all.

It is also noteworthy that the third party may rationally ignore the private information of her ally which, as explained in Section 1, appears to be consistent with claims by the USA that its allies (i.e. Pakistan, Iraq, and Afghanistan) are not doing enough to rein in its enemies. And, as explained in Section 1, the results of this paper have implications for econometric work that attempt to estimate the effect of third-party interventions on conflicts.

It should be obvious from the analysis that the sequential moves in period 1 and the fact that faction A could perfectly observe faction B's choice simplified the third-party inference problem. The problem would have been a lot harder if factions A and B moved simultaneously in period 1.

ACKNOWLEDGMENTS

I thank the editor and two anonymous referees for helpful comments. I also thank Rene Kirkegaard, Asha Sadanand, Dana Sisak, Ruqu Wang, and Jun Zhang for very helpful comments on an earlier version of this paper. My thanks are also due to Olivier Bos, Bram Cadsby, and Osman Ouattara. This paper is a substantially revised version of a paper that was previously titled 'Third party intervention in conflicts and the indirect Samaritan's dilemma,' circulated as CESifo working paper #2711, and presented at the Royal Military College (Canada), the 2011 CESifo ESP conference in Munich, and the 2011 workshop on contests and tournaments at the Technical University of Berlin.

References

Akcinaroglu, S. and Radziszewski, E. (2005) Expectations, rivalries, and civil war duration. *International Interactions* **31** 349–374.

Amegashie, J.A. (2010) On third-party intervention in conflicts: An economist's view. *Peace Economics, Peace Science, and Public Policy* **16**, Article 11.

Amegashie, J.A. and Kutsoati, E. (2007) (Non)intervention in intra-state conflicts. *European Journal of Political Economy* **23** 754–767.

Amegashie, J.A. and Runkel, M. (2012) The paradox of revenge in conflicts. *Journal of Conflict Resolution* **56** 313–330.

Bagwell, K. and Ramey, G. (1991) Oligopoly limit pricing. *Rand Journal of Economics* **22** 155–172.

Balch-Lindsay, D., Enterline, A.J. and Joyce, K.A. (2008) Third-party intervention and the civil war process. *Journal of Peace Research* **45** 345–363.

Betts, R.K. (1996) The delusion of impartial intervention. In *Managing Global Chaos: Sources of and Responses to International Conflict,* edited by C.A. Cocker and F.O. Hampson. Washington, DC: United States Institute of Peace Press, 333–334.

Blattman, C. and Miguel, E. (2010) Civil war. *Journal of Economic Literature* **48** 3–57.

Carment, D. and Harvey, F. (2000) *Using Force to Prevent Ethnic Violence: An Evaluation of Theory and Evidence*. Westport, CT: Praeger.

Carment, D. and Rowlands, D. (1998) Three's company: Evaluating third-party intervention in intrastate conflict. *Journal of Conflict Resolution* **42** 572–599.

Chang, Y.-M. and Sanders, S. (2009) Raising the cost of rebellion: the role of third-party intervention in intrastate conflict. *Defence and Peace Economics* **20** 149–169.

Chang, Y.-M., Potter, J. and Sanders, S. (2007) War and peace: Third-party intervention in conflict. *European Journal of Political Economy* **23** 954–974.

Cho, I.-K. and Kreps, D. (1987) Signaling games and stable equilibria. *Quarterly Journal of Economics* **102** 179–221.

Clark, D.J. and Riis, C. (1998) Competition over more than one prize. *American Economic Review* **88** 276–289.

Collier, P. and Hoeffler, A. (2007) Civil war. In *Handbook of Defense Economics,* edited by K. Hartley and T. Sandler. North Holland: Elsevier, 711–739.

Crawford, T.W. (2005) Moral hazard, intervention and internal war: A conceptual analysis. *Ethnopolitics* **4** 175–193.

Cunningham, D.E. (2010) Blocking resolution: How external states can prolong civil wars. *Journal of Peace Research* **47** 115–127.

Daughety, A.F. and Reinganum, J.F. (2007) Competition and confidentiality: Signaling quality in a duopoly when there is universal private information. *Games and Economic Behavior* **58** 94–120.

Diehl, P., Reifschneider, F.J. and Hensel, P.R. (1996) United Nations intervention and recurring conflict. *International Organization* **50** 683–700.

Dixon, W.J. (1996) Third-party techniques for preventing conflict escalation and promoting peaceful settlement. *International Organization* **50** 653–681.

Elbadawi, I., and Sambanis, N. (2000). External Intervention and the Duration of Civil Wars. World Bank Policy Research Paper 2433. Development Research Group, The World Bank, Washington, DC.

Epstein, G.S. and Gang, I.N. (2009) Good governance and good aid allocation. *Journal of Development Economics* **89** 12–18.

Favretto, K. (2009) Should peacemakers take sides? Major power mediation, coercion, and bias. *American Political Science Review* **103** 248–263.

Fearon, J.D. (1994) Domestic political audiences and the escalation of international disputes. *American Political Science Review* **88** 577–592.

Fearon, J.D. (1995) Rationalist explanations for war. *International Organization* **49** 379–414.

Gent, S.E. (2008) Going in when it counts: Military intervention and the outcome of civil conflicts. *International Studies Quarterly* **52** 713–735.

Gershenson, D. and Grossman, H.I. (2000) Civil conflict: Ended or never ending? *Journal of Conflict Resolution* **44** 807–821.

Grigorian, A. (2005) Third-party intervention and escalation in Kosovo: Does moral hazard explain it? *Ethnopolitics* **4** 195–213.

Grossman, H.I. and Kim, M. (1995) A theory of the security of claims to property. *Journal of Political Economy* **103** 1275–1288.

Harvey, F. (1998) Deterrence failure and ethnic conflict: The case of Bosnia. In *Peace in the Midst of Wars: Preventing and Managing International Ethnic Conflicts,* edited by D. Carment and P. James. Columbia, SC: University of South Carolina Press, 230–264.

Heger, L. and Salehyan, I. (2007) Ruthless rulers: Coalition size and the severity of civil conflict. *International Studies Quarterly* **51** 385–403.

Hertzendorf, M.N. and Overgaard, P.B. (2001) Price competition and advertising signals. *Journal of Economics and Management Strategy* **10** 621–662.

Konrad, K.A. (2009) *Strategy and Dynamics in Contests*. New York: Oxford University Press.

Kuperman, A. (1996) The other lesson of Rwanda: Mediators sometimes do more damage than good. *SAIS Review* **16** 221–240.

Lacina, B. (2006) Explaining the severity of civil wars. *Journal of Conflict Resolution* **50** 276–289.

Leininger, W. (1993) More efficient rent-seeking – a Munchhausen solution. *Public Choice* **75** 43–62.

Morgan, J. (2003) Sequential contests. *Public Choice* **116** 1–18.

Regan, P. (1996) Conditions for successful third party intervention in intrastate conflicts. *Journal of Conflict Resolution* **40** 336–359.
Regan, P.M. (2000) *Civil Wars and Foreign Powers: Outside Intervention in Intrastate Conflict*. Ann Arbor, MI: University of Michigan Press.
Regan, P.M. (2002) Third-party intervention and the duration of intrastate conflicts. *Journal of Conflict Resolution* **46** 55–73.
Regan, P.M. and Stam, A. (2000) In the nick of time: Conflict management, mediation timing, and the duration of interstate disputes. *International Studies Quarterly* **42** 239–260.
Rowlands, D. and Carment, D. (1998) Moral hazard and conflict intervention. In *The Political Economy of War and Peace,* edited by M. Wolfson. Norwell, MA: Kluwer Academic, 267–285.
Rowlands, D. and Carment, D. (2006) Force and bias: Towards a predictive model of effective third-party intervention. *Defence and Peace Economics* **17** 435–456.
Schultz, C. (1996) Polarization and inefficient policies. *Review of Economic Studies* **63** 331–344.
Siqueira, K. (2003) Conflict and third-party intervention. *Defence and Peace Economics* **14** 389–400.
Wagner, R.H. (2007) *War and the State: The Theory of International Politics*. Ann Arbor, MI: University of Michigan Press.
Walter, B.F. (2009) Bargaining failures and civil war. *Annual Review of Political Science* **12** 243–261.
Walter, B.F. and Snyder, J. (1999) *Civil Wars, Insecurity, and Intervention*. New York: Columbia University Press.
Watkins, M. and Winters, K. (1997) Intervenors with interests and power. *Negotiation Journal* **13** 119–142.
Werner, S. (2000) Deterring intervention: The stakes of war and third-party involvement. *American Journal of Political Science* **44** 720–732.
Yuen, A. (2009) Target concessions in the shadow of intervention. *Journal of Conflict Resolution* **53** 745–773.

Appendix A

Recall that $G^*_{k|\mathrm{B}}$ is the investment of faction B of type k in period 1 in the benchmark case of full information, $k=\mathrm{H}$, L. In period 1, I assume that the third party cannot give military assistance. Finally, recall that faction A signals in period 1 if and only if he best responds to faction B's investment in period 1 by choosing an investment which is different from what he (i.e. A) would have chosen if were only maximizing his payoff in period 1. That is, if $G_\mathrm{A} \neq \max\left[0, \sqrt{VG_\mathrm{B}} - G_\mathrm{B}\right]$. Otherwise, faction A is not signaling.

Proof of lemma 1

According to lemma 1, faction A does not signal in a separating equilibrium. So, given that faction B chooses $\hat{G}_{\mathrm{H|B}} > G^*_{\mathrm{H|B}}$ when he is strong and $G^*_{\mathrm{L|B}}$ when he is weak, faction A will choose $\hat{G}_{\mathrm{H|A}} = \sqrt{V\hat{G}_{\mathrm{H|B}}} - \hat{G}_{\mathrm{H|B}} > 0$ and $G^*_{\mathrm{L|A}} = \sqrt{VG^*_{\mathrm{L|B}}} - G^*_{\mathrm{L|B}} > 0$.

Standard refinement criteria (e.g. Cho-Kreps intuitive criterion) for signaling games were developed for games with a single sender. In our game, both factions have the same information which the third party does not have. This is similar to the signaling game in Schultz (1996) where two political parties (the senders) have private information about the cost of providing a public good that voters (the receivers) do not have or Bagwell and Ramey (1991) in which two incumbents (the senders) in a market have private information about cost which a potential entrant (the receiver) does not have.[20] In dealing with his two-sender signaling game, Schultz (1996) found a way of turning a two-sender game into a single-sender game enabling him to adapt the Cho-Kreps intuitive criterion (see also Daughety and Reinganum, 2007). I follow a similar approach.

In particular, when the third party observes any pair, (G_B, $G_\mathrm{A} = \max\left[0, \sqrt{VG_\mathrm{B}} - G_\mathrm{B}\right]$), she infers that faction A is not sending a signal. And so, in this case, the third party can credibly say that he will infer faction B's type based on *only* faction B's action. Accordingly, I specify a *benchmark* set of out-of-equilibrium beliefs for the third party. This *benchmark* specifies the third party's out-of-equilibrium beliefs conditional on *no* signaling from faction A. *If faction A does not signal*, let the third party's out-of-equilibrium beliefs be:

[20]See also Hertzendorf and Overgaard (2001) and Daughety and Reinganum (2007) where two incumbent firms, with possibly different costs, send signals to consumers. A difference between all these papers and mine is that the two informed parties move simultaneously while in my case they move sequentially.

$$\mu(G_B, G_A) = \begin{cases} 1 & \text{if} \quad G_B \in [\hat{G}_{H|B}, \infty) \\ 0 & \text{if} \quad G_B \in [0, \hat{G}_{H|B}) \end{cases}, \tag{A.1}$$

where $G_A = \max\left[0, \sqrt{VG_B} - G_B\right]$.[21] The set of beliefs in (A.1) is common knowledge.

Now, if the third party observes G_B and $G_A \neq \max\left[0, \sqrt{VG_B} - G_B\right]$, what should she infer? Like Schultz (1996), I follow a logic that is similar to the logic of the intuitive criterion. Faction A knows that if he does not signal, the third party will follow the beliefs in (A.1). If $G_B \in [0, \hat{G}_{H|B})$, faction A will not signal because, given (A.1), the third party will believe that faction B is weak even if faction A does not signal. Hence, given (10b) and (12), faction A's equilibrium payoff in period 2 is at its *maximum*. Therefore, given (A.1), not signaling strictly dominates signaling for faction A if $G_B \in [0, \hat{G}_{H|B})$. Accordingly, faction A has the incentive to signal *if and only if* $G_B \in [\hat{G}_{H|B}, \infty)$ because, given (A.1), if he does not signal, the third party will believe that faction B is strong. But then, by the logic of the intuitive criterion, the third party should ignore faction A's signal and believe that faction B is strong whenever she observes $G_B \in [\hat{G}_{H|B}, \infty)$ and $G_A \neq \max\left[0, \sqrt{VG_B} - G_B\right]$. Hence, regardless of faction A's response to G_B, the third party will follow the beliefs in (A.1) and so faction A will not signal. That is, the third party will infer faction B's type from only faction B's actions.

Faction B's payoff over the two periods is

$$\Theta_k(G_B, \mu(G_B)) = \hat{\Pi}_k^1(G_B) + \hat{\Pi}_k^2(\hat{M}(\mu)), \tag{A.2}$$

k = H, L.

In what follows, let $\hat{G}_{H|B}$ be implicitly defined by

$$\Theta_L(G^*_{L|B}, 0) = \Theta_L(\hat{G}_{H|B}, 1). \tag{A.3}$$

Noting that faction A does not signal (i.e. chooses G_A by only maximizing his period-1 payoff), we can write faction B's payoff in period 1 as

$$\hat{\Pi}_k^1(G_B) = \frac{G_B}{G_B + \sqrt{VG_B} - G_B} W_k - G_B = W_k\sqrt{\frac{G_B}{V}} - G_B, \tag{A.4}$$

for k = L, H.

Given that $\hat{\Pi}_L^1$ is decreasing in G_B for $G_B \in [G^*_{L|B}, \infty)$ and $\hat{\Pi}_L^2(\hat{M}(1))$ is independent of G_B, it follows that $\Theta_L(G_B, 1)$ is decreasing in G_B for $G_B \in [G^*_{L|B}, \infty)$.

Then, given $\Theta_L(G^*_{L|B}, 0) < \Theta_L(G^*_{L|B}, 1)$, there exists a unique value of $\hat{G}_{H|B}$ that solves Equation (A.3) *and* satisfies $\hat{G}_{H|B} > G^*_{L|B}$.

I assume that $\hat{G}_{H|B} > G^*_{H|B}$. This holds if $\hat{M}(0)$ is sufficiently bigger than $\hat{M}(1)$. This is because, in this case, the gain to the weak type of pretending to be strong – which is directly

[21]Notice that in considering deviations by faction B, I allow faction A to respond to the deviations. If the third party were not in the game, it would not matter whether I allow faction A to respond or if faction A sticks to his subgame perfect equilibrium investment. In either case, faction B will still choose his subgame perfect equilibrium investment. In this model, because deviations by faction B affect the third party's beliefs, which may be different from his equilibrium beliefs, allowing faction A to respond to deviations by faction B may result in a different choice by faction B than what he will choose if faction A stuck to his equilibrium choice. And since deviations are not observed in equilibrium, we are, of course, asking the following (hypothetical) question: 'if faction B could deviate from equilibrium, will he, mindful of the fact that faction A chooses his investment after observing his (i.e. B) investment, choose an investment that is different from the equilibrium investment?' For consistency, we have to assume that if faction B deviates, he deviates as a *first mover* whose actions are observed by faction A. In contrast, being a *second mover*, faction A does not have to worry about faction B responding to his (i.e. A) deviations.

proportional to the difference $\hat{M}(0) - \hat{M}(1)$ – is so high that it requires a high value of $\hat{G}_{H|B}$ to make him indifferent between his equilibrium payoff and deviating to $\hat{G}_{H|B}$ to obtain $\Theta_L(\hat{G}_{H|B}, 1)$. It is easy to show that

$$\hat{M}(0) - \hat{M}(1) = \left(\frac{\theta SV}{2}\right)^{1/(1-\theta)} \left(\left(\frac{1}{W_L}\right)^{1/(1-\theta)} - \left(\frac{1}{W_H}\right)^{1/(1-\theta)}\right). \quad (A.5)$$

The magnitude of (A.5) depends on the size of the difference between W_H and W_L and the sizes of S and V.

Given (A.1), a separating equilibrium exists if

$$\Theta_H(\hat{G}_{H|B}, 1) \geq \Theta_H(G_B, \mu(G_B)), \quad (A.6)$$

and

$$\Theta_L(G^*_{L|B}, 0) \geq \Theta_L(G_B, \mu(G_B)) \quad (A.7)$$

for all G_B.

The weak type of faction B has no incentive to reduce his investment below $G^*_{L|B}$ because, given the beliefs in (A.1), he reduces his payoff in period 1 without changing his payoff in period 2. Given (A.3) and (A.1), the weak type of faction B has no incentive to increase his investment beyond $G^*_{L|B}$. Therefore, the inequality in (A.7) holds. The strong type of faction B has no incentive to increase his investment beyond $\hat{G}_{H|B}$ because, given the beliefs in (A.1), this reduces his payoff in period 1 without increasing his payoff in period 2. And, I assume that reducing his investment below $\hat{G}_{H|B}$ is not profitable given that this will cause the third party's belief that he is strong to discontinuously fall from 1 to 0. That is, $\Theta_H(\hat{G}_{H|B}, 1) \geq \Theta_H(G_B, 0)$ for $G_B < \hat{G}_{H|B}$.[22] Under these conditions, the inequality in (A.6) holds.

I need to argue that this equilibrium is supported by reasonable out-of-equilibrium beliefs. Because of previous arguments that the third party will infer faction B's type by using only faction B's actions, I can easily adapt the Cho-Kreps 'intuitive criterion' to place restrictions on out-of-equilibrium beliefs in this signaling game. The 'intuitive criterion' requires that out-of-equilibrium beliefs put no weight on types that have no incentive to deviate from a given equilibrium no matter what the third party would conclude from observing the deviation. Given that the payoff of each type of faction B is strictly increasing in μ, it follows that faction B's equilibrium action dominates any out-of-equilibrium action if his equilibrium payoff is higher than any out-of-equilibrium payoff even if such an out-of-equilibrium action causes the third party to believe that faction B is a strong type (i.e. $\mu = 1$). Accordingly, the separating equilibrium above satisfies the Cho-Kreps 'intuitive criterion' if there is no G_B such that

$$\Theta_H(G_B, 1) > \Theta_H(\hat{G}_{H|B}, 1) \text{ and } \Theta_L(G_B, 1) < \Theta_L(G^*_{L|B}, 0), \quad (A.8)$$

and

$$\Theta_H(G_B, 1) < \Theta_H(\hat{G}_{H|B}, 1) \text{ and } \Theta_L(G_B, 1) > \Theta_L(G^*_{L|B}, 0). \quad (A.9)$$

In a separating equilibrium, the strong type enjoys the most advantageous belief (i.e. $\mu = 1$) by the third party so, given that $\hat{G}_{H|B} > G^*_{H|B}$, he has no incentive to deviate to an investment level that is greater than $\hat{G}_{H|B}$. Hence, any deviation must be to a lower level of investment.

[22]The numerical example below satisfies this condition with strict inequality. This was verified by plotting $f(G_B) \equiv \Theta_H(\hat{G}_{H|B}, 1) - \Theta_H(G_B, 0)$ on the domain $[0, \hat{G}_{H|B}]$.

Note that the strong type of faction B will find it profitable to deviate to some $G_{H|B} \in [G^*_{H|B}, \hat{G}_{H|B})$ if the third party will still believe that he is strong. Now, since $\Theta_L(G^*_{L|B}, 0) = \Theta_L(\hat{G}_{H|B}, 1)$ as given in (A.3) and given that $\Theta_L(G_B, 1)$ is decreasing in G_B for $G_B \in [G^*_{H|B}, \infty)$, it follows that $\Theta_L(G^*_{L|B}, 0) < \Theta_L(G_{L|B}, 1)$ for $G_{L|B} \in [G^*_{H|B}, \hat{G}_{H|B})$. Hence, if the strong type of faction B finds a deviation to $G_{H|B} \in [G^*_{H|B}, \hat{G}_{H|B})$ profitable, then the weak type also will find it profitable. Therefore, the beliefs satisfy the Cho-Kreps intuitive criterion for this set of deviations. Now, consider the deviations $G_{H|B} \in [0, G^*_{H|B})$ by the strong type of faction B. Note, from (A.4), that $\partial\hat{\Pi}^1_H/\partial G_B > \partial\hat{\Pi}^1_L/\partial G_B$ for any given G_B. Therefore, even if the third party still holds the most favorable belief (i.e. $\mu = 1$) for faction B, a reduction of investment in period 1 below $G^*_{H|B}$ hurts the strong type of faction B than it hurts the weak type and, in some cases, it is even beneficial to the weak type because $G^*_{L|B} < G^*_{H|B}$. Therefore, it is reasonable for the third party to believe that such reductions are by the weak type as in Equation (13).[23] Hence, the beliefs in (A.1) satisfy the intuitive criterion.

Like any signaling game, this one has multiple equilibria. In particular, there are separating equilibria in which the strong type of faction B chooses $G_{H|B} > \hat{G}_{H|B}$ and the weak type chooses $G^*_{L|B}$. However, for any $G_{H|B} > \hat{G}_{H|B} > G^*_{H|B}$, the strong type of faction B is strictly better off by deviating to $\hat{G}_{H|B}$ if the third party will believe that he is strong. And given (A.3), this deviation is not profitable for the weak type even if the third party believes that he is strong. Therefore, it is reasonable for the third party to believe that such deviations are by the strong type.[24] Therefore, the strong type of faction B will deviate from any separating equilibrium where $G_{H|B} > \hat{G}_{H|B}$.

Now given (A.3) and the fact that the weak type does not signal in a separating equilibrium, there cannot be a separating equilibrium with $G_{H|B} < \hat{G}_{H|B}$. Hence, the equilibrium in lemma 1 is the unique separating equilibrium. **QED**

Example: Suppose that $S = 8$, $W_H = 4$, $W_L = 3.2$, $V = 3$, and $\theta = 0.5$. Then, $G^*_{H|B} = 0.25 (W_H)^2/V = 1.333$ and Equation (A.3) gives $\hat{G}_{H|B} = 1.9842 > G^*_{H|B}$. $\Theta_H(\hat{G}_{H|B}, 1) = 2.034, \Theta_L(G^*_{L|B}, 0) = 0.900, \hat{M}(0) = 2.516, \hat{M}(1) = 1.250, G^*_{L|B} = 0.853, G^*_{L|A} = 0.746$ and $\hat{G}_{H|A} = 0.455$. Also, $\Theta_H(\hat{G}_{H|B}, 1) \geq \Theta_H(G_B, 0)$ for $G_B < \hat{G}_{H|B}$. Given the third party's military assistance, all equilibrium investment levels and payoffs in period 2 are positive.

Appendix B

Proposition 3: *There is no pooling equilibrium which satisfies the Cho-Kreps intuitive criterion.*

To prove proposition 3, I proceed in two steps. I begin by proving the following lemma:

Lemma 3(a): *An equilibrium in which faction B pools and faction A signals does not exist.*

Proof: As before, faction A signals in period 1 if and only if he best responds to faction B's investment in period 1 by choosing an investment which is different from what he (i.e. A)

[23]In this case, the intuitive criterion actually does not tell us what to do. However, the reasoning used here is in the spirit of the D1 condition in Cho and Kreps (1987). The D1 condition requires that we put the *entire* weight on the type that is willing to deviate for a wider range of inferences by the receiver (i.e. uninformed party). In my case, while a decrease (deviation) that the strong type of faction B finds profitable is also profitable to the weak type, the converse is not true. That is, there are some decreases in investment that the weak type finds profitable but are not profitable to the strong type. Hence, it is reasonable for the third party to set $\mu = 0$ in (A.1) for investments smaller than $\hat{G}_{H|B}$. This is why this kind of reasoning is in the spirit of the D1 condition of Cho and Kreps (1987).

[24]In this case, the inequalities in (A.8) are applicable with the inequality for the weak type being a weak inequality.

would have chosen if were only maximizing his payoff in period 1. That is, if $G_A \neq \max\left[0, \sqrt{VG_B} - G_B\right]$.

Consider a pooling equilibrium in period 1 where (a) both types of faction B choose $G_{L|B} = G_{H|B} = \hat{G}_B$ and (b) faction A chooses $\tilde{G}_A \neq \hat{G}_A = \max\left[0, \sqrt{V\hat{G}_B} - \hat{G}_B\right]$. Given a pooling equilibrium, the third party's equilibrium belief should be $\mu = q \in (0,1)$.

Regardless of B's type, suppose faction A deviates to $\hat{G}_A$. Then, he is not signaling, so the third party should stick to the belief $\mu = q$. Then, this deviation increases faction A's payoff in period 1 without changing his payoff in period 2. Therefore, faction A will deviate from the pooling equilibrium. **QED**

Next, I prove the following lemma:

Lemma 3(b): *There is no equilibrium satisfying the intuitive criterion in which both types of faction B pool but faction A does not signal.*

Proof: As before, by faction A does not signal, I mean that he best responds to faction B's investment in period 1 by ignoring his payoff in period 2.

Suppose there exists an equilibrium in which both types of faction B pool in period 1, faction A does not signal, and the equilibrium satisfies the intuitive criterion. That is, suppose that (a) both types of faction B choose $G_{L|B} = G_{H|B} = \hat{G}_B$ and (b) faction A chooses $\hat{G}_A = \max\left[0, \sqrt{V\hat{G}_B} - \hat{G}_B\right]$. Given a pooling equilibrium, the third party's equilibrium belief should be $\mu = q \in (0,1)$.

Given that faction A is the first mover in period 2, his equilibrium payoff in this period is

$$\hat{\Pi}^2_{k|A} = \frac{V^2\left(1 + \hat{M}(q)\right)^{\theta}}{4W_k}, \tag{B.1}$$

$k =$ H, L and his equilibrium payoff in period 1 is

$$\hat{\Pi}^1_A = \frac{\hat{G}_A}{\hat{G}_A + \hat{G}_B}V - \hat{G}_A. \tag{B.2}$$

Note, from Equations (B.1) and (B.2), that faction A's payoff in period 1 is independent of faction B's type while his payoff in period 2 depends on faction B's type. Given that $\hat{G}_A$ is faction A's best response to $\hat{G}_B$ based on only his period-1 payoff, it follows that faction A's payoff in period 1 is decreasing in his investment for any $G_A \neq \hat{G}_A$. It suffices to consider only deviations where $G_A > \hat{G}_A$. Suppose that for any such increase in investment, the third party holds the most favorable belief that benefits faction A (i.e. $\mu = 0$), so that faction A's payoff in period 2 increases because the third party military assistance increases discontinuously from $\hat{M}(q)$ to $\hat{M}(0)$. Given (B.1) and $W_H > W_L$, we know that

$$\frac{\partial \hat{\Pi}^2_{L|A}}{\partial \hat{M}} > \frac{\partial \hat{\Pi}^2_{H|A}}{\partial \hat{M}}. \tag{B.3}$$

Therefore, while from Equation (B.2), an increase in investment beyond $\hat{G}_A$ reduces faction A's expected payoff by the same amount in period 1 regardless of faction B's type, the inequality in (B.3) shows that it increases his payoff in period 2 by a bigger amount if faction B is weak than if faction B is strong. A *marginal* increase in investment beyond $\hat{G}_A$ has only a negligible (marginal) reduction in payoffs in period 1 but yields a discontinuous increase in payoffs in period 2 if such a deviation will lead to the most favorable belief (i.e. μ) in faction A's interest. Then, there exists some $G_A > \hat{G}_A$ at which faction A is indifferent between deviation and nondeviation if faction B is strong but is strictly better off by deviating to this level of investment if faction B is weak. Hence, if the third party observes this investment in period 1, it is reasonable for her, in the spirit of the intuitive criterion, to believe that this deviation was by a faction A whose opponent was the weak type of faction B. Hence, if his opponent is the weak type of faction B, then faction A will deviate from a pooling equilibrium in which he does not signal. Therefore, the equilibrium does not satisfy the intuitive criterion.

Lemmas 3(a) and 3(b) imply proposition 3. **QED**